Washington Assessment of Student Learning

Washington Alternate Assessment System (WAAS)

2002

Technical Report

Prepared by
The Riverside Publishing Company

for
Office of the Superintendent of Public Instruction
P.O. Box 47220
Olympia, Washington 98504-7220

May 2003

Part 1: Overview and Background

Introduction

The *Washington Alternate Assessment System (WAAS)* was administered operationally for the second time during the Spring 2002. In the fall of 2002, four levels of performance relative to the grade level academic content and achievement standards on the WAAS assessments were established. The *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 1999) recommends that test developers and publishers produce a technical manual that provides information documenting the technical quality of an assessment, including evidence for the reliability and validity of test scores. This document contains the technical information for the 2002 WAAS.

State assessment programs provide one method of determining student academic achievement. The Washington State Assessment System provides accountability for program and educational opportunities for all students. Alternate assessment, as part of Washington's assessment program, ensures a unified system, program, and student accountability linked to the common core of learning within the general curriculum.

The Washington Alternate Assessment System (WAAS) process was developed by the Washington Alternate Assessment Task Force (Appendix A) and expanded by Advisory Panels (Appendix b and C) in response to the following requirement in the Individuals with Disabilities Education Act 1997: "The State has established goals for the performance of children with disabilities in the state that . . . are consistent, to the maximum extent appropriate, with other goals and standards for children established by the state." It was toward fulfillment of this requirement that alternate assessments are based on Washington's Essential Academic Learning Requirements (EALRs) in the content areas of Communication, Reading, Writing, and Mathematics (Appendix D). In this manner, all students in Washington will be moving toward the same general standards. The inclusion of students with disabilities in the assessment and accountability system is critical to ensure appropriate allocation of resources and learning opportunities for these students.

The Washington Alternate Assessment System was designed for a very small percentage of the total school population for the *Washington Assessment of Student Learning (WASL)*, even with accommodations, would be an inappropriate measure of progress (Appendix E). The two options currently available in the alternate assessment system are commercially available tests and portfolio assessment.

Purpose of the Portfolio Assessment

The Washington Alternate Assessment Task force, made up of administrators, higher education personnel, teachers, and parents, determined the following two-fold purpose of the portfolio assessment:

- To provide an appropriate method of measuring progress on state goals and standards for students who are not able to access the WASL or any commercially available test, even with accommodations and
- To ensure that students will be able to generalize the Individualized Education Program (IEP) skills to the maximum extent possible.

The basic building block of the portfolio assessment is evidence of the student's performance and progress toward reaching IEP goals. Each of the entries in the portfolio document two dimensions of learning: progress on IEP skills linked to the EALRs and student generalization of those skills.

Deleted: ev

Portfolio evidence should demonstrate participation in and progress toward those IEP goals that are aligned to state standards (EALRs). In this way, evidence of progress on IEP skills linked to the EALRs can measure progress on state goals and standards.

Portfolio evidence should also show the extent to which a student can demonstrate and generalize the IEP skill linked to EALRs in the following ways:

- using appropriate modifications/adaptations, supports, or assistive technology in order to demonstrate all he or she knows and is able to do;
- in a variety of settings and contexts in which the student is able to use learned skills. These places can include the classroom, other areas of the school, community settings, and home;

- interacting with nondisabled peers and others during IEP activities for the purpose of developing social relationships to enrich his or her life; and
- using self-determination skills in planning, monitoring and evaluating IEP skill activities.

Purpose of the Commercially Available Tests

The Individualized Education Program (IEP) team may select a commercially available test (Appendix F) to measure progress toward state standards in listening, reading, writing, or mathematics. This option was available for students whose academic skills can be measured, but whose disability prevents them from participating in one or more content areas of the WASL, even with accommodations. A commercially available test (CAT) should only be administered in content areas for which the student qualifies for specially designed instruction.

For the 2002 administration of commercially available tests, a list of acceptable tests was provided. IEP teams chose from the list an appropriate test for the student that measured the student's skills in the content area. No alignment to specific standards was required.

Participation Rates

The participation by district varied (Table 1). District participation depended on the district's approach to providing programming and assessment for student with special needs. Some districts did not submit any commercially available scores but did submit a number of portfolios, other districts submitted a number of portfolios but did not report any commercially available scores. Some district did not report either commercially available test results or portfolios.

Table 1: Number of Districts Participating

	Washington Assessment of Student Learning	Commercially Available Test Only	Portfolio Submitted
Number of Districts	296	152	84

Federal guidance letters indicate that states should develop alternate assessment participation guidelines so that approximately 1-2% of the student population is eligible for an alternate assessment in each given year. One would expect between 2,100 and 4,200 of the students in grade 4, 7 and 10 to be assessed using either a commercially available test or portfolio. As can be seen in Table 2 the number of portfolios submitted together with the number of commercially available tests submitted is about in the this estimate. About 15% of the students in the WAAS were assessed using the Portfolio Assessment while about 85% of the student had results from a Commercially Available Test reported.

Table 2: Number of Students Assessed in Grades 4, 7 and 10 by Type of Assessment

	Total Number of Students											
	Washington Assessment of Student Learning				Commercially Available Test Only				Portfolio Submitted			
Subject	Gr4	Gr7	Gr10	Total	Gr4	Gr7	Gr10	Total	Gr4	Gr7	Gr10	Total
Listening	75,310	75,846	70,272	221,428	971	625	409	2,005	171	140	116	427
Reading	75,222	75,420	69,960	220,602	1,397	969	704	3,070	171	138	116	425
Writing	74,503	74,476	68,341	217,320	1,287	944	701	2,932	171	140	116	427
Math	75,431	75,511	70,128	221,070	1,212	898	689	2,799	170	140	117	427

Use of Commercially Available Tests

Table 3 indicates the frequency of commercially available tests that were used.

Table 3: Frequency of Use of Commercially Available Test by Subject

	Frequency			
Name of Instrument	**Listening**	**Reading**	**Writing**	**Mathematics**
Woodcock Johnson (III)	557	756	769	708
Brigance	579	651	634	650
Woodcock Johnson (WJ-R)		858	802	719
Wechsler Individual Achievement Test WIAT	371	318	306	274
ITBS	136	150	130	144
Mini-Battery of Achievement		125	124	117
Wide Range Achievement Test	28	90	79	88
Peabody Picture Vocabulary Test	109			
Oral Written Language Scales	63		7	
Kaufman Test of Educational Achievement		19	16	20
Peabody Individual Achievement Test		15	16	13
Key Math				40
Mini-Battery of Achievement	33			
Stanford Achievement Test, 9th Ed. – Hearing Impaired		8	6	8
Clinical Evaluation of Language Fundamentals	19			
Diagnostic Achievement Battery	7	2	5	5
Test of Auditory Comprehension of Language	14			
Woodcock Reading Mastery Test		14		
Test of Math Ability 2				11
Test of Written Language			11	
Test of Language Development	6			
Diagnostic Achievement Test for Adolescents	4			
OTHER	65	182	139	169

Part 2: Scoring

Commercially Available Scoring

School personnel administered and scored the commercially available tests according to the publishers' instructions. The scorer entered these scores onto the demographic sheet for WAAS (Appendix G). When applicable, the portfolio scores were also entered onto this sheet. That sheet was then scanned and the reports were generated from that information.

Portfolio Scoring

The portfolios were scored over a two-week period in June. For the first week, a small group of teachers and representatives from the Riverside Publishing Company (RPC) and Pearson Educational Measurement (PEM) were led by OSPI in range-finding. Teachers RPC and PEM personnel were trained by OSPI so that they all had a common understanding of scoring dimension definitions and score points for each dimension in the portfolio.

There are five scoring dimensions divided into two parts, with one dimension scored on specific content area sections of the portfolio and four dimensions scored across the entire portfolio. Part I scores for Progress on IEP skills are determined based on evidence in separate portfolio entries for Communication, Reading, Writing, and Mathematics. Part II scores for Student Generalization of Skills in four dimensions are determined by examining evidence across the entire portfolio. The content area Part I score is added to the total of the four dimension scores in Part II to obtain a Total Score for the content area. Thus, four separate total scores are generated for the student (one total score for each content area).

OSPI staff pre-selected a number of portfolios that exemplified score points for each dimension. First, two of the aforementioned portfolios were used as tools to train teachers, RPC and PEM staff. Teachers, RPC and PEM personnel were given one

portfolio to score. When all were finished scoring, OSPI discussed each score given and consensus was achieved. This step was repeated three times.

Once teachers, RPC and PEM staff were trained to OSPI's standards, the group was divided into three groups of two teachers and one RPC or PEM staff. Each group scored portfolios and then all groups met to come to consensus. Appendix H shows the score sheet used. Fourteen portfolios were scored in this manner. RPC, PEM, and OSPI personnel reviewed all of the scored portfolios and selected four portfolios to be used in training teacher scorers. Scoring summaries and annotations were written to accompany the training sets for the Portfolio Scoring Institute.

The second week, additional teachers were used as scorers. The teachers who had attended the first week served as table leaders and the Riverside Publishing Company and Pearson Educational Measurement staff served as assistant table leaders. The first day was used as a full day of training, and scoring started on the second day. Teachers were trained by OSPI so that they all had a common understanding of dimension definitions and score points for each dimension in the portfolio. OSPI led the training on definitions of each dimension and its rubric. OSPI pre-selected two portfolios that exemplified all ranges of score points for each dimension. OSPI facilitated discussion of these portfolios. Teachers were given two portfolios to score independently. OSPI and RPC facilitated discussion upon completion of scoring. When OSPI and RPC concluded that all teachers were properly trained, scoring procedures were reviewed.

On the first day every portfolio was scored twice and the table leader (or assistant table leader) score was used as the final score. When clarification was needed, or discrepancies were found, OSPI staff served as the final arbiter.

Four tables were established for scoring purposes. At each table there was a table leader (RPC person or teacher returning from range-finding) and four teacher scorers. There was a lead scorer (OSPI) table, as well. Scorer reliability was calculated at this table. When clarification was needed, or discrepancies found, OSPI was the final arbiter.

Scorers chose a portfolio randomly. Portfolios were arranged according to district. Scorers were told not to choose a portfolio from their district or their table leader's district. Scorers signed for one portfolio with its unique number. At each table was a sheet on which scorers were required to check in and out with their initials. Next, scorers scored the portfolios and then recorded their scores on content and dimension sheets. Scorers gave portfolio and paper work to table leader.

Table leaders initialed and scored each portfolio without looking at teacher scorers' results. Table leaders scored portfolio "blind." Table leaders filled out an entry form with each student's name and portfolio number. Table leaders filled in and transferred all scores onto bubble sheet. Table leaders handed each portfolio to lead scorer's table and scores were entered into a database.

Part 3: Validity and Reliability of the Commercially Available Tests

As noted a range of commercially developed tests was used. Each commercially available test publishes evidence for the validity of that test. The Special Education Assessment Advisory Panel for standard setting purposes the percentile rankings from each of the tests are comparable to the standard score regardless of the testing instrument.

Part 4: Reliability of the Portfolio Assessment

Introduction

The reliability of assessment scores is a measure of the degree to which the scores on the test are a "true" measure of the examinees' knowledge and skill relevant to the tested knowledge and skills. There are several ways to obtain estimates of score reliability: test-retest, alternate forms, internal consistency, and generalizability analysis are the most common. Test-retest estimates require administration of the same instrument at different times. In a sense a portfolio system is a collection of evidence from a full school year and as such should increase the reliability of the measurement of a student's ability.

However, no evidence was collected to confirm this speculation. Alternate forms reliability estimates require administration of two parallel assessments. These tests must be created in such a way that we have confidence that they measure the same domain of knowledge and skills using different items. Unfortunately at this time there is only one set of evidence collected

The scoring design for the 2002 assessment did not readily allow for estimating the rater variance component. However, inter-score agreement and coefficient alpha were two internal consistency measures used to estimate score reliability.

Inter-Scorer Agreement

Inter-scorer agreement is an important source of evidence for the reliability of test scores. When two trained judges agree with the score given to a student's work, this gives support for the score on the short-answer or extended response item. To determine the degree to which judges gave equivalent scores to the same student work the percent of agreement between scorers was examined. Reliability of scoring was determined by looking at the difference between the score from the teacher scorers and the table leader scorers (Table 4). Although these percentages of agreement would appear to be reasonable, as can be seen in table 6 the approximately 60 percent of the portfolios were given a score of 1 on each dimension. Since most scores were the same one would expect agreement among scorers.

Table 4: Percentage Agreement Between Teacher and Table Leader Scores

Grade 4	**Percentage**		
Amount of Agreement	**Day 1**	**Day 2**	**Day 3**
Scores exact the same	64.3%	70.4%	61.2%
Scores are different by 1	30.9%	27.5%	29.4%
Scores are different by 2	4.1%	2.1%	7.4%
Scores are different by 3	0.7%	0.0%	2.0%

Table 4: Percentage Agreement Between Teacher and Table Leader Scores (cont)

Grade 7			
Amount of Agreement	**Day 1**	**Day 2**	**Day 3**
Scores exact the same	61.2%	66.3%	66.7%
Scores are different by 1	26.8%	30.6%	25.7%
Scores are different by 2	8.2%	1.6%	6.3%
Scores are different by 3	3.8%	1.6%	1.4%
Grade 10			
Amount of Agreement	**Day 1**	**Day 2**	**Day 3**
Scores exact the same	70.6%	75.7%	61.4%
Scores are different by 1	24.6%	16.9%	27.2%
Scores are different by 2	4.4%	5.2%	9.9%
Scores are different by 3	0.4%	2.2%	1.5%

Periodically during each day, and at the end of each day, the scores reliabilities for each teacher and for the group were calculated by each dimension/trait. On day 2, 3 and 4 portfolios scored by teachers who had 70% or above exact agreement received one score. However, every third portfolio was still double-scored by a table leader or assistant table leader to check on reliability. In these cases the table leader scores were still used as the final score. If teachers fell below 70% exact agreement, the portfolios they scored were double-scored on every portfolio. When reliabilities were low by trait, then the scorer was re-trained.

In a traditional large-scale assessment, there is a chance that a student's handwriting or neatness could influence a scorer's judgment of the portfolio evidence. In the case of the portfolios, the skill, training and ability of the person putting the portfolio together could influence the scorer's impression and score. As teachers gain more experience in this activity, it is likely that differences in portfolios from the ability of the teacher to select materials for the portfolio will even out, and it will be possible to ensure that the scores received are not unduly influenced by the presentation of the portfolio. For 2003, it has been recommended that a new score be available for scoring that indicates that there is not enough evidence in the portfolio to score.

Coefficient Alpha

Coefficient Alpha is a score reliability index of internal scale consistency/homogeneity. Alpha can be estimated from scores obtained on one occasion and is appropriate when a score is intended to measure a single trait. Table 5 provides the Coefficient Alpha for the Total scores and Part II scores. As indicated in the associated formula, the value of Alpha is affected by the number of components making up a score, the variance of the individual components, and the total score variance. In the context of the WAAS Total scores and Part II scores, relatively higher values of Alpha will tend to result when the total scores have greater variability and/or the scores across the individual components are very similar (i.e., internally consistent). This reliability index is only sensitive to random errors associated with this source of score variability. It does not incorporate temporal errors (as would a test-retest reliability index) or random error associated with rater variance (addressed elsewhere in this document). Systematic sources of variance, such as rater effects, might artificially increase these values.

$$\text{Alpha} = (N/N\text{-}1) * (1 - \Sigma\text{Var(part)}/\text{Var(total)})$$

Where N = Number of components combined to form total
ΣVar(part) = Sum of the variance for the individual components
ΣVar(total) = Variance of the total scores

Table 5: Coefficient Alpha for Total Scores and Part II Scores

	Total Score				
	Communication	Reading	Writing	Math	Part II
Grade 4	0.89	0.87	0.88	0.89	0.90
Grade 7	0.89	0.90	0.90	0.90	0.91
Grade 10	0.83	0.84	0.83	0.82	0.83

Part 5: Description of Performance of Students

Table 6 provides a summary of the percentage of students obtaining each of the scale scores in each dimension that was scored plus the mean and standard deviation for each dimension. As last year the majority of students were given a scale score of 1 for most

dimensions but there appears to be a reduction in the percent being given a score of 1 in 2002. There tended to be more scores of 0 awarded this year compared to 2001. As well there tended to be more scores of 2, 3, and 4 in 2002 than in 2001. In 2001, at grade 10, no students were awarded a score of 4 on six of the eight dimensions while in 2002 only the dimension of Self Determination had 0% of students obtaining a 4 while 5.2% of the portfolios in Mathematics and 7.6% on Modifications were scored at level 4. As in 2001, the scores in 2002 for Modifications tend to be higher than for the other dimensions. The scores awarded for Modifications tend to be more spread out than for the other dimensions as well.

Standards were established in the fall of 2002. For a description of the standard setting process and the results by standard please see part 6.

Table 6: Percentage of Students Obtaining Each Score on the Portfolio and Average Score By Grade

Grade 4

		Part I				Part II			
		Communication	**Reading**	**Writing**	**Math**	**Modifications**	**Settings**	**Social Relations**	**Self Determination**
Percentage of Students Obtaining Each Score	**0**	8.2%	9.1%	9.4%	9.1%	5.9%	5.9%	5.9%	5.9%
	1	52.5%	51.9%	57.2%	56.5%	38.5%	50.9%	60.9%	63.9%
	2	25.9%	30.5%	21.4%	22.7%	24.9%	23.1%	23.7%	20.1%
	3	11.4%	3.9%	8.2%	9.7%	18.3%	14.8%	8.3%	5.9%
	4	1.9%	4.5%	3.8%	1.9%	12.4%	5.3%	1.2%	4.1%
Number of Portfolios Scored		158	154	159	154	169	169	169	169

Grade 7

		Part I				Part II			
		Communication	**Reading**	**Writing**	**Math**	**Modifications**	**Settings**	**Social Relations**	**Self Determination**
Percentage of Students Obtaining Each Score	**0**	19.6%	19.9%	22.4%	20.1%	17.8%	17.8%	17.8%	17.8%
	1	54.1%	45.7%	48.1%	51.7%	33.8%	42.0%	56.1%	59.2%
	2	13.5%	20.5%	16.0%	15.4%	24.2%	24.8%	21.0%	15.3%
	3	8.1%	8.6%	9.0%	10.7%	17.2%	13.4%	3.8%	5.1%
	4	4.7%	5.3%	4.5%	2.0%	7.0%	1.9%	1.3%	2.5%
Number of Portfolios Scored		148	151	156	149	157	157	157	157

Grade 10

		Part I				Part II			
		Communication	**Reading**	**Writing**	**Math**	**Modifications**	**Settings**	**Social Relations**	**Self Determination**
Percentage of Students Obtaining Each Score	**0**	12.2%	7.8%	7.8%	6.0%	5.0%	5.0%	5.0%	5.0%
	1	55.7%	60.9%	67.0%	64.7%	44.5%	53.3%	72.5%	69.2%
	2	20.9%	19.1%	20.9%	19.8%	31.9%	28.3%	17.5%	21.7%
	3	7.8%	7.8%	2.6%	4.3%	10.9%	11.7%	4.2%	4.2%
	4	3.5%	4.3%	1.7%	5.2%	7.6%	1.7%	0.8%	0.0%
Number of Portfolios Scored		115	115	115	116	119	120	120	120

PART 6 – STANDARD SETTING

Introduction

The Federal legislation and regulations for ESEA and IDEA reauthorization requires states to report results for all students assessed using general assessments and alternate assessments relative to the same grade level academic content and achievement standards. In anticipation of the federal government publication of a Notice for Proposed Rule Making to allow setting alternate achievement standards for students with the most significant cognitive disabilities who participate in alternate assessments, the Office of Superintendent of Public Instruction sought to establish four levels of performance based on alternate achievement standards on the WAAS assessments in the fall of 2002.

Development Of Standard Setting Procedures And Review Of Results

The Special Education Assessment Advisory Panel met on June 5, 2002, to begin discussions on setting standards on alternate assessments so that student performances could be reported. The panel reviewed the standard setting methodology used by three states, and reviewed synthesis reports and policy directives from the National Center for Educational Outcomes (NCEO). The advisory panel also reviewed the types of scores generated for commercially available tests and portfolio alternate assessments.

In 2002, the Commercially Available Test (CAT) option of the Washington Alternate Assessment System (WAAS) allowed districts to select from a list of 24 tests (table 3) from which student results could be reported.

Dr. Thomas Hirsch, under contract with Riverside Publishing Company, worked with Dr. Marty McCall and Nancy Arnold of the Office of Superintendent of Public Instruction to develop a set of procedures for standard setting for the WAAS portfolios. The Washington National Technical Advisory Committee (TAC) reviewed and made recommendations relative to these procedures at its October 21-22, 2002 meeting. The Washington Special Education Assessment Advisory Panel reviewed these procedures at

its November 21, 2002 meeting. In addition, at this meeting the advisory panel accepted the proposed standard setting methodology for the Commercially Available Test option.

To set standards on the WAAS portfolio assessment, a variation on a holistic method of standard setting that Riverside Publishing has used in the past was recommended to the panel and the TAC. The method is loosely based on Jaeger and Mills' method (Cizek, 2001). With this method, the panelists' task is to classify student work into one of several performance categories defined to capture levels of performance as expressed by the performance-level categories. The method is holistic in that the panelists consider the whole of an individual student's work. The panelists review folders of student portfolios sampled to represent the full range of scores, and are asked to sort these folders into four performance levels as represented by the quality of the students' work.

With this WAAS portfolio standard setting method, panelists would review the implications of their standards in the form of impact data. Panelists would receive cumulative frequency distributions of student scores that allow them to see the percent and number of students in each category given the standards the group of panelists has set. These data would be made available to panelists after they have completed the two sorting processes.

The procedures use these standard-setting methodologies and a consideration of standard-setting principles in order to optimize the efficacy of this process. The goal of the standard setting is to recommend performance thresholds or cut scores for the WAAS portfolio assessment that have been established by Washington educators, subject matter experts, and administrators in the best interests of students and the overall educational process. For each subject matter, the final achievement category is decided from a minimum subject matter rating and a minimum combination of ratings on generalization skills. The object of the standard setting process is to decide how to combine the ratings so that the portfolios can be placed in four achievement categories. The object is not to come up with a new rubric or to change student scores but to find a way to combine them so that these students can be included in reports and evaluations.

For the portfolio option, a standard setting meeting was held January 15-16, 2003. Dr. Hirsch was the principal facilitator of the standard setting process. Nancy Arnold as the OSPI Alternate Assessment Specialist explained the ESEA and IDEA background that is the impetus for standards, the history of the WAAS portfolio, and presented the academic achievement standard descriptions. Dr. McCall documented the process to ensure adherence with the planned standard setting steps.

The results of standard setting for both the commercially available tests and portfolios were shared with the Special Education Assessment Advisory Panel at its March 18, 2003 meeting and with the National Technical Advisory Committee at its April 14, 2003 meeting. The results of standard setting were then presented to the Superintendent of Public Instruction.

Procedure for Setting Standards for the Commercially Available Test Option

The methodology for setting standards was to be straightforward so that results expressed in standard scores, age equivalents or grade equivalents could be grouped. The Special Education Assessment Advisory Panel members first created levels of performance by percentile rankings since these numbers are comparable to most standardized scores. There were 24 testing instruments allowed for the Commercially Available Test option. Three types of scores were generated by these tests: standard scores (based on IRT scales), age equivalent, or grade equivalent scores. The advisory panel also recommended that age equivalent scores or grade equivalent scores be used only if standard scores were not reported. The panel determined three cuts for percentiles corresponding with standard scores so that results could be reported in four achievement levels.

A small subcommittee of the advisory panel recommended corresponding cut scores for age equivalent scores and grade equivalent scores. Additional rules were set: age equivalent scores were to be used if no standard scores were reported and grade equivalent scores would used only if no standard scores or age equivalent scores were

reported. Standard scores would be used to establish levels of performance if multiple types of scores were reported. If no scores were reported, they would be reported as "no data" and would be considered as students not meeting the standard in accountability calculations.

The Special Education Assessment Advisory Panel also understood throughout this standard setting process that the federal Office of Elementary and Secondary Education had disallowed the Commercially Available Test option for future use as a state alternate assessment. This option was disallowed because these tests were not aligned with the EALRs and were not comparable between different test instruments. The panel understood that standards would be set for the sole purpose of reporting assessment results to the public as required by IDEA and ESEA.

Procedures for Setting Standards on the Portfolio

Selection of Standard Setting Panelists

Table 7 indicates the composition of the 14-member standard setting panel that was selected to set the standards for the Portfolio. Each person participating in the standard-setting process was selected for his or her qualifications as a judge of student performance based on various factors. Teachers, educators, and subject area experts who were selected as panelists exemplified the required subject-area knowledge, teaching experience, and/or understanding of students necessary for an appropriate and comprehensive standard-setting study. Each panelist participating in the process represented the knowledge and understanding of his or her peers throughout the course of the process, lending a balance between diverse opinion and consensus. To ensure that a diverse opinion was obtained, some of the participants had not participated in making recommendations during the development of the portfolio system or in the scoring of the portfolios.

Table 7: List of Portfolio Standard Setting Panel

Participant	Job Title	Previous Involvement in WAAS	District
Sue Alfawicki	Teacher		Bellingham SD
Jessica Dadisman	Teacher	Scorer	Orting SD
Dalrae Danilson	Teacher	Scorer	Mount Vernon SD
Charlene Esget	Teacher		Bellingham SD
Gordon Fischer	Principal, School Board Member		Tukwila SD
Julie Fleisch	Teacher	Scorer	Spokane SD
Jeannette Forman	Teacher	Scorer, Advisory Panel	Longview SD
Mike Jacobsen	Special Education Director	Advisory Panel	White River SD
Carol Johnson	Special Education Director	Advisory Panel	Richland SD
Keith Mars	Special Education Director	Advisory Panel	Fife SD
Sherry Mashburn	Parent		
Julie Moore	Teacher		Central Kitsap SD
Rachel Quenemoen	Senior Fellow for Technical Assistance and Research	Advisory Panel	National Center for Educational Outcomes (NCEO)
Bev Sweet	Parent	Advisory Panel	

To ensure balance, a stratified sample of school district staff and other stakeholders throughout Washington were contacted. A concerted effort was made to balance the panel based on county representation, urban representation, representation of students with significant disabilities and schools serving various sizes of populations, gender and race/ethnicity. The overarching goal of consensus in this forum is not the unanimous agreement of all parties, but the bringing together of individual divergent experiences to form a common understanding of student performance in a subject area that is truly larger, and broader, than its individual parts.

Initial Procedure

Following the approval of the standard setting methodology, portfolios with adequate evidence were identified by the Office of Superintendent of Public Instruction Alternate

Assessment Specialist. Forty scored portfolios with sufficient evidence had been copied for range finding activities for 2001 and 2002. These portfolios had also been used as training sets during training of portfolio scorers. The total scores on the portfolio content entries ranged from 5 to 20 points. OSPI staff recorded the scores for each of five scoring dimensions in each content area of the portfolios. Sixty-two scoring patterns were found. Frequencies of student score profiles were generated from this group. Fourteen portfolios were selected to be used for standard setting activities; this subgroup represented 32 different scoring patterns. Standard setting was conducted using only these portfolios with adequate evidence.

Panel Meeting - Step 1 - Training

The purpose of this step was to give information about what the task of standard setting entails and about how student results will be used in reporting and evaluation. Participants introduced themselves and provided some information about their backgrounds to help the panelists get to know each other and provide a perspective for the various backgrounds of the panelists. The process and criteria for selecting panelists was reviewed to explain why the panelists were there and provide insight into what factors are important to the standard-setting process. Panelists were told that their job was to recommend how to place student results in four levels of performance. Background on the ESEA and IDEA requirements was provided.

A presentation that describes the standard-setting process was made that focused on the general nature of standard setting. This helped the panelists understand the overall process and the iterative nature of the standard setting. This session did not focus on specific procedures that the panelists will use later, but attempted to give them an appreciation for the group judgment process and the panelist's role in the process.

The agenda for the two days was reviewed to give the panelists a good perspective on the process and the pace of the meeting. A general question and answer period was held. General administrative tasks such as filling in of expense claims and signing of security forms were completed at the end of this step.

Panel Meeting - Step 2 - Review of the Assessment Material

Panelists became familiar with the assessment at this point. Panelists were told that the portfolio is a collection of evidence of student work and that only the evidence that is present can be considered. The panelists reviewed the participation criteria for WAAS so that they were familiar with the type of student who would participate in the portfolio assessment (Appendix E).

Panelists were introduced to the content validity evidence for the assessment and the scoring processes. Exemplars of student work for each score point were reviewed by the panelists in a group setting so questions about scoring and rubrics could be answered. A summary of the development process was provided. Panelists were informed as to the use of standards across assessments and issues of comparability. The Portfolio Scoring Summary (Appendix I) and the EALR Extensions (Appendix F) were distributed to panelists (Appendix D). Both the Scoring dimensions and scoring criteria and EALR Extensions were developed and refined by a special education stakeholder curriculum group during the piloting phase of portfolio development. Both of these documents were explained in detail to panelists.

In 2003, evidence for student performance in science will be collected for the first time. Given the nature of the standard setting procedures and that the requirements for portfolio entries are similar for each content area, standards were set for all content areas including science. Panelist raised questions related to scoring and were reminded that the current task before the committee is standard setting and not to alter scoring procedures.

Panel Meeting - Step 3 - Understanding the Definitions of the Standards

This step was designed to introduce panelists to the definitions of the standards. Panelists did a brainstorming exercise to help them think of students and student work that typify the definitions of each standard and the performance of students who are at the standard. Panelists did not write or re-write the definitions at any time. This step only served to familiarize panelists with definitions that have previously been determined, and to help the panelists think about students who are at each of the levels.

Academic Achievement Standard Descriptions were provided to the panelists that identify student performance in four levels (Appendix J). These descriptions were written by the Alternate Assessment Specialist, using NAEP descriptors and the portfolio scoring criteria as a framework. The basis for these definitions was reviewed, and the panelists discussed the definitions until the levels or categories of student performance were clearly distinguished from each other and no ambiguity regarding their characteristics remained.

Panelists were broken into three groups of 5 or 6 each and a record of these suggestions was made so that they could be discussed by the entire group and used as a reference during the standard-setting process. Each subgroup was composed of at least one person who had portfolio-scoring experience, at least one person on the Advisory Committee, and others who were new to the portfolio process. Panel subgroups discussed and revised the descriptors. The three subgroups shared their revisions with the entire panel.

Panel Meeting - Step 4 - Determination of Achievement Levels Based on Scoring Guides and Scoring Patterns

Panelists received rating sheets, and were instructed in the process of completing the sheets. Working with the table of frequencies of scoring patterns and with the existing descriptors for each scale, panelists decided on the category ("meets" or "not meets" proficiency) for each score pattern. The level placement for each pattern and the minimum total score in each category (level 2, 3, 4) were recorded. This set of ratings was placed under Round 1 Levels on the rating forms (Appendix K).

Panel Meeting - Step 5 - Holistic Classification of a Range of Student Portfolios

Participants received feedback on the overall panel cut scores they had established. The cut scores for the panel were based on the mean of the minimum total scores for each proficiency category. The OSPI Alternate Assessment Specialist compiled and selected 32 scoring patterns found in 14 portfolios with sufficient evidence from the 2001 and 2002 WAAS administration.

Panelists were then given the set of 14 scored portfolios that were exemplars of the patterns already categorized. Working individually, they reviewed the portfolio evidence, total score, and scoring patterns. They made decisions for each portfolio and reset, if needed, the minimum score for each proficiency level. They then reconvened into subgroups (same members as those on the previous day) to discuss ratings and make final decisions. The discussion at this point included whether or not to set a minimum score for the Part I score for each content area. Each panel member completed the rating form for Round 2.

Panel Meeting - Step 6 - Review of Impact Data

Panelists were presented impact data in the form of frequencies for each score pattern generated from 2002 portfolio scores. They were provided with statewide performance data to judge the impact of group standards - these data included the entire population of students assessed that satisfied state-determined completion criteria. They received frequency distributions of total scores for the state and scoring patterns. Panelists discussed the impact of standards on the state.

Participants discussed the impact data and resumed the discussion of a minimum Part I score. After reviewing the Academic Achievement Standard Descriptions, the panel agreed on a method for determining standards.

Panelists were given their Final Rating Sheets and asked to make any changes they wish on the basis of the impact data, and group discussions. Panelists were advised that this is the last round of adjustments. Panelists were allowed to change the raw score value of the group cut scores according to this new information. No group consensus was pursued. Participants turned in their final Round 3 cut score recommendations. The final cut scores were calculated as the mean of the minimum total scores for each proficiency category.

Panel Meeting - Step 7 - Evaluation of the Standard-Setting Process

Panelists were given evaluation forms to complete and open-ended comments were encouraged. A summary of the responses is given in Table 8. The eight evaluation

questions were rated on a 1 to 5 scale, with 1 indicating *Not At All* to 5 indicating *Completely*. All of the questions had a majority of responses of 4 or 5. Six panelists included additional comments. Four of the six comments were positive, such as great two days, great process, and good interaction. One panelist made a suggestion for training the panelists on the general education curriculum and related (WASL) assessment procedures. One panelist felt that some bias was present (men seemed to be listened to more so than women.)

Table 8: Percentage and Number of Standard Setting Panel Members Agreeing with Statements Regarding Process*

	Not At All				***Completely***
Question	1	2	3	4	5
1. Understand scoring process	0%	7% (1)	0%	50% (7)	43% (6)
2. Understand descriptions of performance for each level	0%	7% (1)	7% (1)	43% (6)	43% (6)
3. Descriptions accurately reflect performance levels	0%	7% (1)	7% (1)	57% (8)	29% (4)
4. Comfortable with rating the scoring patterns	0%	7% (1)	14% (2)	29% (4)	50% (7)
5. Comfortable with ratings after viewing portfolios	0%	14% (2)	7% (1)	21% (3)	57% (8)
6. Balance of content (Part I) vs. Generalization (Part II) parts of scoring patterns	0%	14% (2)	14% (2)	29% (4)	43% (6)
7. Agree with final cut scores proposed by panel	0%	0%	7% (1)	14% (2)	79% (11)
8. Comfortable with process used to set standards	0%	0%	7% (1)	21% (3)	71% (10)

***N=14**

RESULTS

Commercially Available Tests

The Special Education Assessment Advisory Panel determined three cut scores for percentiles corresponding with standard scores so that results could be reported in four achievement levels, as follows:

Level 1 – missing or inaccurate results (such as standard scores above or below corresponding percentile)
Level 2 – Scores below the 25th percentile

Level 3 – Scores on or between the 25th and 50th percentile

Level 4- Scores above the 50th percentile

Table 9 shows the score ranges for the corresponding cut scores for standard scores, age equivalent scores or grade equivalent scores. Standard scores will be used to set the performance level. Age equivalent scores will be used only when standard scores are not provided. Grade equivalent scores will only be used if no standard scores or age equivalent scores are provided.

Table 9: Commercially Available Tests with Levels, Standard Scores and Equivalent Scores

Levels	Percentile Range	Standard Score (Grades 4, 7, 10)	Age Equivalent Grade 4	Age Equivalent Grade 7	Age Equivalent Grade 10	Grade Equivalent Grade 4	Grade Equivalent Grade 7	Grade Equivalent Grade 10
Level 4	Above 50th percentile	101 to 135	Greater than 9.0	Greater than 12.0	Greater than 15.0	Greater than 4.8	Greater than 7.8	Greater than 10.8
Level 3	On or between 25th & 50th percentile	90 to 100	9.0	12.0	15.0	4.8	7.8	10.8
Level 2	Below 25th percentile	66 to 89	Less than 9.0	Less than 12.0	Less than 15.0	Less than 4.8	Less than 7.8	Less than 10.8
Level 1	Missing or inaccurate scores	Below 66 or over 135	NA	NA	NA	Other scores	Other scores	Other scores

For reporting and accountability purposes, student results reported in Level 3 or 4 are considered as meeting the standard. Student results reported in Level 1 or 2 are counted as students not meeting the standard. Students with no data and students, who were not tested, even though their Individualized Education Program (IEP) specifies alternate assessments, will be included as students not meeting the standard. Students whose parents refuse to allow their child (opt out) to participate in the Commercially Available Test assessment option will also be counted similarly to a student not tested. One of the weaknesses of the CAT option is the difficulty in comparing scores from different testing instruments. Even though most students were assessed using only three test instruments (Woodcock-Johnson - Revised, Brigance, or Wechsler Individual Achievement Test), the

scores between these instruments are not comparable. This alternate assessment option will not be available in future years due to this issue.

Portfolio

Alternate achievement standards were set for the alternate assessment portfolio. Cut scores were set using the total score from adding the score for each part one score (Part 1 score - Progress on IEP Skills score for the content area) to the total score for the part two dimension (Modifications and Adaptations, Settings and Contexts, Social Relationships, and Self-Determination). Table 10 shows the mean score from the standard setting committee for round 1 and round 2.

Table 10: Portfolio Cut score Recommendations by Round

	Round 1	**Round 2**
Level 2	8	8
Level 3	11	12
Level 4	16	16

The second part of the decision rule requires a minimum score on the first scoring dimension (Progress on IEP Skill). The achievement level for any subject cannot be more than 1 level higher than the subject dimension score. That is, a portfolio with a 1 in the subject dimension cannot be in an achievement category higher than 2. A portfolio with a score of 2 in the subject dimension cannot be in an achievement category higher than 3. Table 11 summarizes the decision rule. Portfolios with insufficient evidence

Table 11: Decision Rule for Determining Level of Performance on WAAS portfolio

Level	Total Score*	Part I Score Required on Progress on IEP Skill
4	16 to 20	3 or 4
3	12 to 15	2 or 3 or 4
2	8 to 11	1 or 2 or 3 or 4
1	5 to 7	1 or 2 or 3

* Total score = progress in content area (Part I score) + mod + set + soc+ self. Portfolios with insufficient evidence would be reported separately as IE and would not be reported in one of the performance levels.

would be treated as zeros and students would be reported as not meeting the standard for accountability purposes. Appendix L shows the relationship among the various total scores, level, and pattern of various scores.

Tables 12, 13, 14, and 15 show the number and percentage students achieving standards on the Commercially Available Tests and on the Portfolio Assessment for each content area. For students who did not have any information for either the commercially available test or the portfolio were assigned proportionately to portfolio and test groups. The achievement standards reported here were for the WAAS assessments and should not be compared to the results or standards for students taking the WASL.

Table 12: Number and Percentage of Students Achieving Standards on the Commercially Available Tests And on the Portfolio Assessment of the WAAS in 2002 by Grade Mathematics

	Portfolio and Commercial Test Combined Groups				Portfolio			Commercial Test		
	4th	7th	10th		4th	7th	10th	4th	7th	10th
Number of Students Tested:	1,382	1,038	806		170	140	117	1,212	898	689
Percent Who Met Standard*:	**13.9%**	**6.3%**	**4.1%**		**12.9%**	**10.0%**	**4.3%**	**14.0%**	**5.7%**	**4.1%**
Level 4 (exceeding standard):	53	17	12		3	0	1	50	17	11
Level 3 (meeting standard):	139	48	21		19	14	4	120	34	17
Percent Not Meeting Standard*:	**86.1%**	**93.7%**	**95.9%**		**87.1%**	**90.0%**	**95.7%**	**86.0%**	**94.3%**	**95.9%**
Level 2 (below standard):	709	525	437		45	34	27	664	491	410
Level 1 (below standard):	481	448	336		103	92	85	378	356	251
No Data for Math Portfolio or Commercial Test	311	168	107	586	38.3	22.7	15.5	272.7	145.3	91.5
	100.0%	100.0%	100.0%		12.3%	13.5%	14.5%	87.7%	86.5%	85.5%
Total	1,693	1,206	913		208	163	133	1485	1043	780
Percent meeting standard including students w/ no data	**11.3%**	**5.4%**	**3.6%**		**10.6%**	**8.6%**	**3.8%**	**11.4%**	**4.9%**	**3.6%**

Table 13: Number and Percentage of Students Achieving Standards on the Commercially Available Tests And on the Portfolio Assessment of the WAAS in 2002 by Grade Reading

	Portfolio and Commercial Test Combined Groups			Portfolio			Commercial Test		
	4th	7th	10th	4th	7th	10th	4th	7th	10th
Number of Students Tested:	1,568	1,107	820	171	138	116	1,397	969	704
Percent Who Met Standard*:	**9.8%**	**9.5%**	**7.0%**	**12.9%**	**13.0%**	**6.9%**	**9.4%**	**9.0%**	**7.0%**
Level 4 (exceeding standard):	49	30	19	4	1	0	45	29	19
Level 3 (meeting standard):	105	75	38	18	17	8	87	58	30
Percent Not Meeting Standard*:	**90.2%**	**90.5%**	**93.0%**	**87.1%**	**87.0%**	**93.1%**	**90.6%**	**91.0%**	**93.0%**
Level 2 (below standard):	891	602	457	56	45	34	835	557	423
Level 1 (below standard):	523	400	306	93	74	75	430	326	231
No Data for Reading Portfolio or Commercial Test	125	100	93	13.6	12.5	13.2	111.4	87.5	79.8
	100.0%	100.0%	100.0%	10.9%	12.5%	14.1%	89.1%	87.5%	85.9%
Total	1,693	1,207	913	185	150	129	1,508	1,057	784
Percent meeting standard including students w/ no data	**9.1%**	**8.7%**	**6.2%**	**11.9%**	**12.0%**	**6.2%**	**8.8%**	**8.2%**	**6.3%**

Table 14: Number and Percentage of Students Achieving Standards on the Commercially Available Tests And on the Portfolio Assessment of the WAAS in 2002 by Grade Writing

	Portfolio and Commercial Test Combined Groups			Portfolio			Commercial Test		
	4th	7th	10th	4th	7th	10th	4th	7th	10th
Number of Students Tested:	1,458	1,084	817	171	140	116	1,287	944	701
Percent Who Met Standard*:	**8.5%**	**6.5%**	**3.5%**	**13.5%**	**10.0%**	**4.3%**	**7.8%**	**6.0%**	**3.4%**
Level 4 (exceeding standard):	43	28	9	3			40	28	9
Level 3 (meeting standard):	81	43	20	20	14	5	61	29	15
Percent Not Meeting Standard*:	**91.5%**	**93.5%**	**96.5%**	**86.5%**	**90.0%**	**95.7%**	**92.2%**	**94.0%**	**96.6%**
Level 2 (below standard):	795	557	456	57	51	34	738	506	422
Level 1 (below standard):	539	456	332	95	75	77	444	381	255
No Data for Writing Portfolio or Commercial Test	235	123	96	27.6	15.9	13.6	207.4	107.1	82.4
(assigned proportionately to portfolio and test groups)	100.0%	100.0%	100.0%	11.7%	12.9%	14.2%	88.3%	87.1%	85.8%
Total	1,693	1,207	913	199	156	130	1,494	1,051	783
Percent meeting standard including students w/ no data	**7.3%**	**5.9%**	**3.2%**	**11.6%**	**9.0%**	**3.9%**	**6.8%**	**5.4%**	**3.1%**

Table 15: Number and Percentage of Students Achieving Standards on the Commercially Available Tests And on the Portfolio Assessment of the WAAS in 2002 by Grade Communication Skills

	Portfolio and Commercial Test Combined Groups			Portfolio			Commercial Test		
	4th	7th	10th	4th	7th	10th	4th	7th	10th
Number of Students Tested:	1,142	765	525	171	140	116	971	625	409
Percent Who Met Standard*:	**15.1%**	**12.9%**	**8.8%**	**14.6%**	**8.6%**	**7.8%**	**15.2%**	**13.9%**	**9.0%**
Level 4 (exceeding standard):	84	33	12	4	1	2	80	32	10
Level 3 (meeting standard):	89	66	34	21	11	7	68	55	27
Percent Not Meeting Standard*:	**84.9%**	**90.5%**	**93.1%**	**85.4%**	**91.4%**	**92.2%**	**84.8%**	**86.1%**	**91.0%**
Level 2 (below standard):	611	383	262	54	49	36	557	334	226
Level 1 (below standard):	358	283	217	90	75	74	268	208	143
No Data for Com. Portfolio or Commercial Test	551	442	388	82.5	80.9	85.7	468.5	361.1	302.3
	100.0%	100.0%	100.0%	15.0%	18.3%	22.1%	85.0%	81.7%	77.9%
Total	1,693	1,207	913	254	221	202	1,439	986	711
Percent meeting standard including students w/ no data	**10.2%**	**8.2%**	**5.0%**	**6.2%**	**4.6%**	**4.2%**	**12.3%**	**10.1%**	**5.7%**

References

Cizek, G.J., Ed. 2001. *Setting Performance Standards: Concepts, Methods, and Perspectives*. Lawrence Erlbaum Associates: Mahwah, NJ.

Appendix A

Alternate Assessment Task Force - 1997 to August 2000

			School District
Ms.	Virginia	Alonzo	Clover Park School District
Ms.	Nancy	Arnold	OSPI
Ms.	Judy	Bean	Colville
Ms.	Sheila	Bell	Central Valley
Mr.	Michael	Cashion	Colville
Ms.	Kathy	Christiansen	ESD 101
Ms.	Marcia	Davidson	Western Washington University
Mr.	Tom	Delaney	
Ms.	Cindy	Egan	Selah School District
Ms.	Linda	Elman	Central Kitsap
Mr.	Ron	Franklin	Green Hill Academic School
Ms.	Faye	Fuchs	ESD 105
Mr.	Forest	Hertlein	Mukilteo
Ms.	Kay	Jakutis	Shoreline
Mr.	Mark	Jewell	Federal Way
Ms.	Debra	Knesal	ESD 114
Mr.	Randy	Lake	Teacher
Mr.	Hans	Landig	Wapato
Ms.	Jeannene	London	North Thurston School District
Mr.	Duncan	MacQuarrie	Tacoma Public Schools
Ms.	Peggy	Mayer-Chelgren	Lake Stevens Middle School
Mr.	Hans	Michielsen	East Valley
Ms.	Darcy	Miller	Washington State University
Ms.	Minnie	Obregon	Wenatchee High School
Ms.	Sandra	Owen	Pullman
Ms.	Abbie	Pack	Richard Gordon Elementary
Ms.	Lois	Parks	Elma
Ms.	Shirley	Ramsey	Tenino-Rainier SD
Ms.	Joan	Seeberger	ESD 113
Mr.	Ron	Sherman	ESD 105
Ms.	Barbara	Tompkins	SEAC
Ms.	Jennifer	Traufler	Wenatchee
Dr.	Gordon	Wallace	Kiona-Benton School District
Mr.	Ric	Williams	Everett School District

Appendix B

Alternate Assessment Curriculum Work Group - February 2000

Work Group Members

- Lynnda Biek - VI teacher
- Nancy Arnold - OSPI Special Education
- Laura Bolt - Reading Teacher
- Teri Nickerson - Special Ed Teacher
- Tammy Droppo - Math Teacher
- Ginger Alonzo - District Special Ed Admin
- Lesley Thompson - OSPI Reading Specialist
- Jeannene London - Special Ed Teacher
- Joan Seeberger -ESD Special Ed Admin
- Fonda Abbey - Special Ed Teacher
- Holly Seifert - District SLP
- Elaine Talbot - Special Education State Needs Project Coordinator

Facilitators (ILSSA):

- Jacqui Farmer Kearns
- Steve Stafford
- Paula Burdette

Advanced Systems

- Julie Armentrout
- Chris Beeso

Appendix C

Alternate Assessment Advisory Panel - November 2000 to present

Dr. Mark Jewell
Curriculum, Instruction and Assessment Director
Federal Way School District

Dr. Gale Hanninen
Director of Special Services
Sumner S. D.

Carla Jackson, Executive Director
Kent School District

Dan Kelly, Director Special Services
West Valley School District

Debra Knesal, Special Education Director
ESD 114, Bremerton

Linda Sullivan-Dudsic, SLP
Bremerton School District

Jeannene London, Teacher
Mark Twain Elem., Pasco School District

Fonda Abbey, Teacher
Evergreen Elem., Clover Park SD

Gail Hasbrouck, School Psychologist
Special Education Services
Yakima School District

Mary O'Leary Christensen
Special Ed Coordinator
Tacoma School District

Carol Johnson
Richland School District

Betsy Minor Reid, Special Serv Coor
North Central ESD 171, Manson SD

Dr. Gary Livingston, Superintendent
Educational Service District 113
Olympia

Lucille Nollette, Asst.
Director/Special Ed
Bellingham School District

Ms. Nancy Skerritt, Assistant Superintendent
Tahoma School District

Rachel Quenemoen, NCEO
National Technical Advisor to States

Mike Jacobsen, Student Support Coordinator
White River School District

Keith Mars, Director of Special Serv
Fife School District

Bev Sweet
FEPP and Parent

Ron Cammaert
Riverside Publishing

OSPI Staff:
Terry Bergeson
Greg Hall
Bob Harmon
Kathy Bartlett

Mary Alice Heuschel
Nancy Arnold
Doug Gill

Appendix D

Essential Academic Learning Requirements (EALR) Extensions For Communication, Reading, Writing, Mathematics, Science

The critical function of the EALRs, the access skills, instructional activities, and assessment strategies contained in this document will assist special education staff members in linking functional IEP skills to the EALRs, in providing access to the general education curriculum, and in measuring student progress toward achieving the EALRs.

What are the basic attributes of communication?

Communication consists of verbal or non-verbal cues or skills that allow the student to gain understanding (receptive) or to impart a message (expressive).

What are the basic attributes of reading?

Reading consists of pictures, symbols, words, and/or text that have meaning and which the reader decodes to construct meaning.

What are the basic attributes of writing?

Writing consists of encoding symbols in a way that results in a product and conveys meaning.

What are the basic attributes of mathematics?

Mathematics consists of a language of symbols, numbers and words that communicates about patterns and relationships that allow the student to participate in mathematical inquiry and problem-solving**.**

What are the basic attributes of science?

Science consists of questioning cause and effect phenomena and using technology to make personal sense of the world and to solve problems.

For the most current version of the EALR extensions document, please link to:
www.k12.wa.us/SpecialEd/pubdocs/EALR_Extension%20_Guide_Oct_02.pdf

Appendix E

Participation Guidelines for the Washington Alternate Assessment System

The purpose of the Washington Alternate Assessment System (WAAS) is to provide appropriate ways to measure progress on EALRs for students in special education programs for whom the WASL is inappropriate, even with accommodations. The decision for a student to participate in the Washington Alternate Assessment System (WAAS) must be based on the unique needs of the individual student, not a specific disability category, time spent in the general education classroom, or program placement. The IEP team must ensure that the decision for a student to participate in the WAAS is **not** solely based on prior knowledge that the student would perform poorly on general state tests; ongoing disruptive behavior; the result of excessive or extended absences or social, cultural, or economic differences.

1. To be eligible for participation in the WAAS, the student must have a current IEP that documents the need for an alternate assessment.

2. To participate in the WAAS, the student must be at the appropriate grade level (4, 7, and 10). Students with no grade level assignment will need to be assessed at least three times during their educational career (approximately at ages 9, 12 and 15).

3. IEP teams may opt to use ***commercially available tests*** to measure progress in reading, math, writing or listening. This option is available for students whose academic skills can be measured, but whose disability prevents them from participating in one or more component parts of the WASL even with accommodations.

4. If an IEP team determines that any component part of the WASL is inappropriate for a student and that commercially available tests are also inappropriate, the ***portfolio*** assessment should be used.

5. The percentage of students in special education programs participating in an alternate assessment (either *commercially available tests* or the *portfolio* assessment) should not exceed 20 percent of the special education population in the district at grades 4, 7, and 10 in a given year.

For further information on the participation of students with disabilities in the state's assessment programs, please see *Guidelines for IEP Teams in Determining WASL Assessment Options for Students in Special Education Programs*, and *Guidelines for Participation and Testing Accommodations for Special Populations on the Washington Assessment of Student Learning (WASL)*, Olympia, WA: Office of Superintendent of Public Instruction. These documents are also available at the following web site: www.k12.wa.us/SpecialEd/assessment.aspx

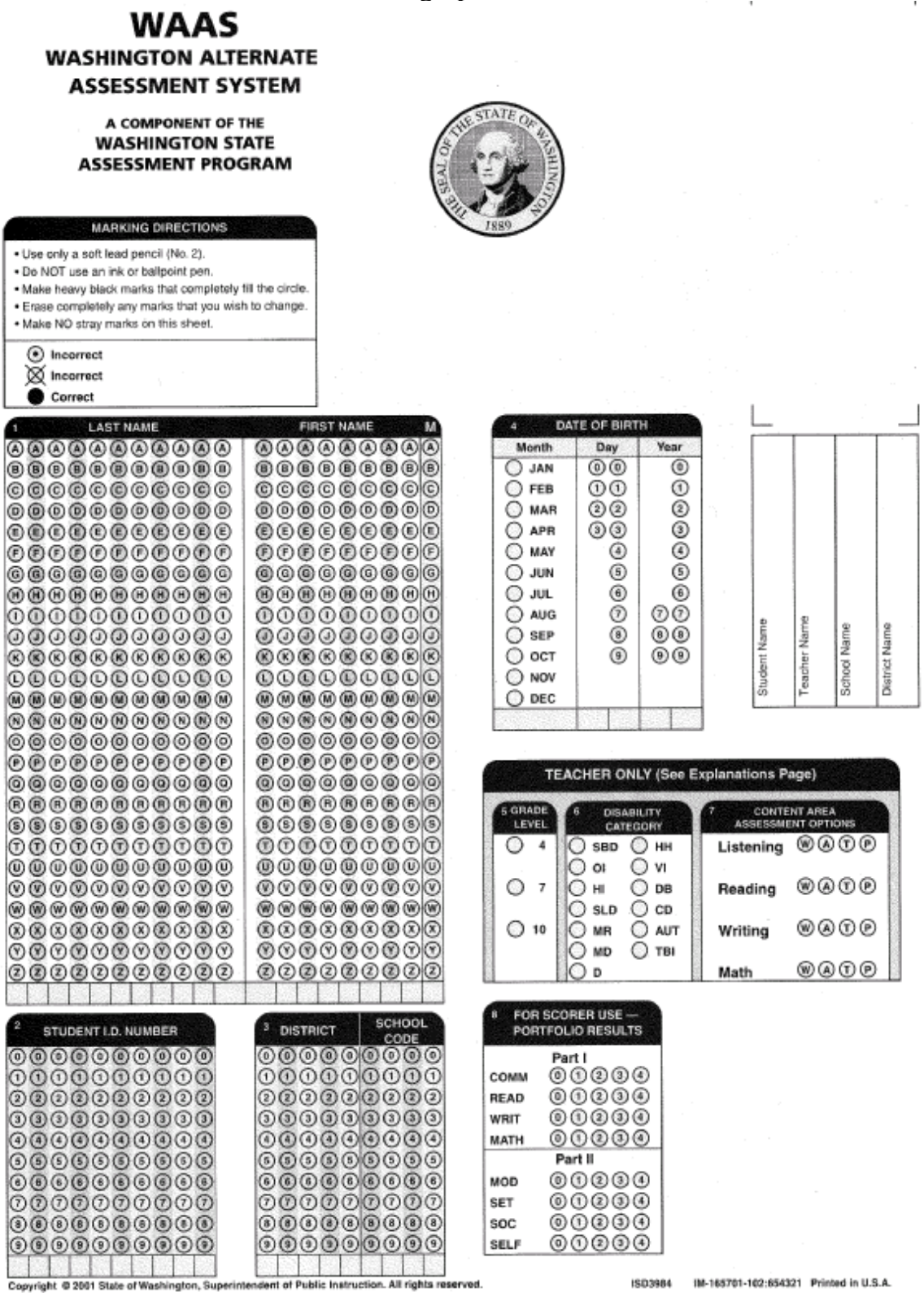
Appendix F

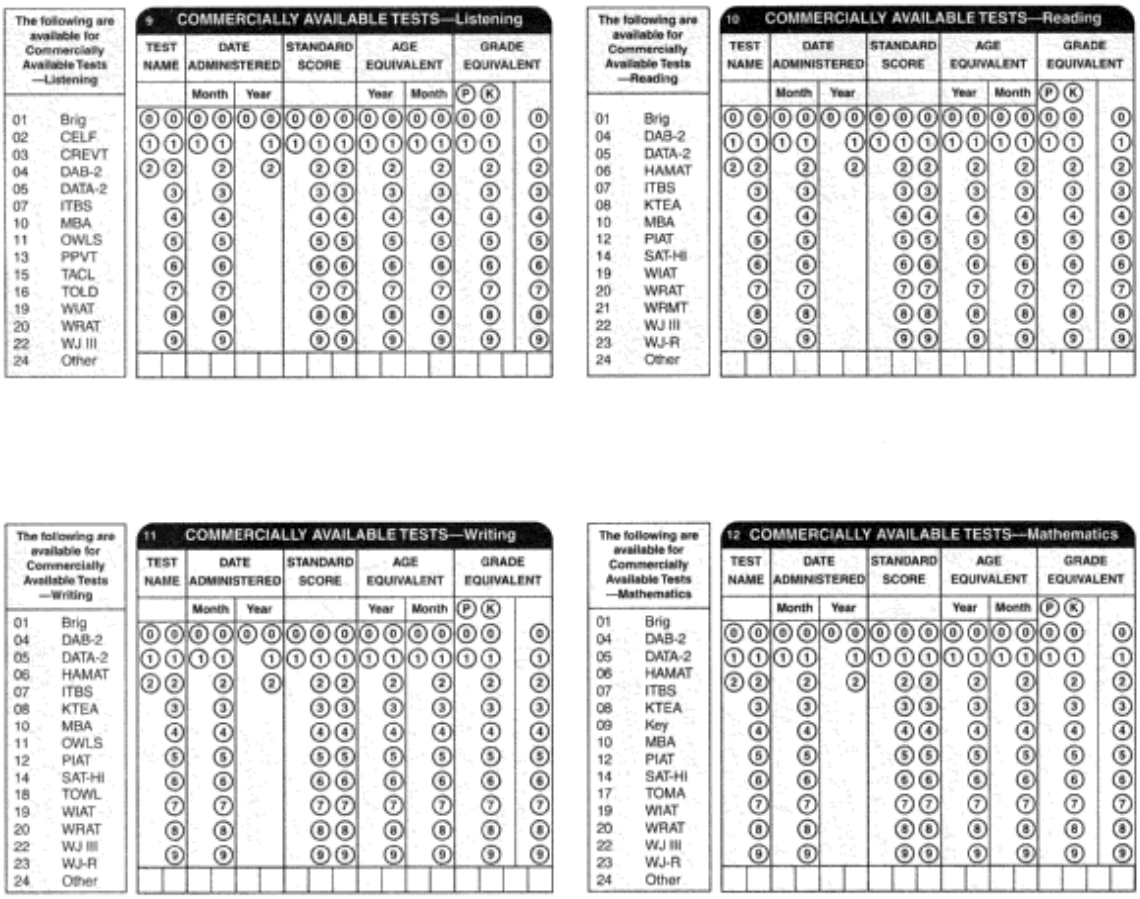
List of Commercially Available Tests Used For WAAS 2002

TEST NAME	CONTENT AREA			
Brigance (Brig)	L	R	W	M
Clinical Evaluation of Language Fundamentals (CELF)	L			
Comprehensive Receptive & Expressive Vocabulary Test (CREVT)	L			
Diagnostic Achievement Battery (DAB-2)	L	R	W	M
Diagnostic Achievement Test for Adolescents (DATA-2)	L	R	W	M
Hammill Multiability Achievement Test (HAMAT)		R	W	M
Iowa Test of Basic Skills (ITBS)	L	R	W	M
Kaufman Test of Educational Achievement (KTEA)		R	W	M
Keymath (Key)				M
Mini-Battery of Achievement (MBA)	L	R	W	M
Oral Written Language Scales (OWLS)	L		W	
Peabody Individual Achievement Test (PIAT)		R	W	M
Peabody Picture Vocabulary Test (PPVT)	L			
Stanford Achievement Test, 9th Ed.- Hearing Impaired (SAT-HI)		R	W	M
Test of Auditory Comprehension of Language (TACL)	L			
Test of Language Development (TOLD)	L			
Test of Math Ability (TOMA)				M
Test of Written Language (TOWL)			W	
Wechsler Individual Achievement Test (WIAT)	L	R	W	M
Wide Range Achievement Test (WRAT)	L	R	W	M
Woodcock Reading Mastery Test (WRMT)		R		
Woodcock-Johnson III (WJ III)	L	R	W	M
Woodcock-Johnson-Revised (WJ-R)		R	W	M
Other Commercially Available Test that meets Criteria (Other)	L	R	W	M

Appendix G

WAAS Washington Alternate Assessment System Demographic Sheet





The following are available for Commercially Available Tests —Listening	
01	Brig
02	CELF
03	CREVT
04	DAB-2
05	DATA-2
07	ITBS
10	MBA
11	OWLS
13	PPVT
15	TACL
16	TOLD
19	WIAT
20	WRAT
22	WJ III
24	Other

9 COMMERCIALLY AVAILABLE TESTS—Listening

TEST NAME	DATE ADMINISTERED		STANDARD SCORE	AGE EQUIVALENT		GRADE EQUIVALENT		
	Month	Year		Year	Month	P	K	

The following are available for Commercially Available Tests —Reading	
01	Brig
04	DAB-2
05	DATA-2
06	HAMAT
07	ITBS
08	KTEA
10	MBA
12	PIAT
14	SAT-HI
19	WIAT
20	WRAT
21	WRMT
22	WJ III
23	WJ-R
24	Other

10 COMMERCIALLY AVAILABLE TESTS—Reading

TEST NAME	DATE ADMINISTERED		STANDARD SCORE	AGE EQUIVALENT		GRADE EQUIVALENT		
	Month	Year		Year	Month	P	K	

The following are available for Commercially Available Tests —Writing	
01	Brig
04	DAB-2
05	DATA-2
06	HAMAT
07	ITBS
08	KTEA
10	MBA
11	OWLS
12	PIAT
14	SAT-HI
18	TOWL
19	WIAT
20	WRAT
22	WJ III
23	WJ-R
24	Other

11 COMMERCIALLY AVAILABLE TESTS—Writing

TEST NAME	DATE ADMINISTERED		STANDARD SCORE	AGE EQUIVALENT		GRADE EQUIVALENT		
	Month	Year		Year	Month	P	K	

The following are available for Commercially Available Tests —Mathematics	
01	Brig
04	DAB-2
05	DATA-2
06	HAMAT
07	ITBS
08	KTEA
09	Key
10	MBA
12	PIAT
14	SAT-HI
17	TOMA
19	WIAT
20	WRAT
22	WJ III
23	WJ-R
24	Other

12 COMMERCIALLY AVAILABLE TESTS—Mathematics

TEST NAME	DATE ADMINISTERED		STANDARD SCORE	AGE EQUIVALENT		GRADE EQUIVALENT		
	Month	Year		Year	Month	P	K	

Submittal Procedures

1. Complete a 2002 WAAS Demographic form for every student in special education programs who is participating in an alternate assessment (commercially available test or portfolio). Do not complete a WAAS form for special education students participating in WASL for all four content areas.

2. Complete the demographic information on the WASL booklets for students participating in alternate assessments. These WASL booklets should be boxed and returned in the same manner as all other scored materials.

3. Insert WAAS forms into the front of portfolios. Collect all WAAS forms for students participating in commercially available tests (do not send the test protocol or score report). Group WAAS demographic forms and portfolios for your school and return WAAS forms and portfolios to your district WAAS contact person.

4. Return your completed WAAS demographic forms and portfolios to your District WAAS Test Coordinator.

Appendix H

*Student Name*_____ *Portfolio Number*_____

Washington Alternate Assessment Portfolio Scoring Summary

Part I: Progress on IEP Skills

(Progress on IEP skills scored for each content area entry.)

	1	2	3	4
Progress on IEP Skill linked to EALRs	Little or no progress on targeted skills linked to the EALRs in portfolio entry.	Clear progress on targeted skills linked to the EALRs in portfolio entry.	Attains goal for targeted IEP skills linked to the EALRs in portfolio entry.	Exceeds goal for targeted IEP skills linked to the EALRs in portfolio entry.

Part II: Student Generalization of Skills

(These dimensions are scored across the entire portfolio.)

	1	2	3	4	Scorer Use Only
Modifications and Adaptations	No or limited evidence that the student uses supports, modifications, adaptations or assistive technology in portfolio entries.	The student appropriately uses supports, modifications, adaptations or assistive technology in some portfolio entries.	The student appropriately uses supports, modifications, adaptations or assistive technology in most portfolio entries.	The student appropriately uses natural supports, modifications, adaptations or assistive technology within and across all portfolio entries.	
Settings and Contexts	Student participates in a limited number of settings or use of targeted skills unclear in portfolio entries.	Student performs targeted skills in some settings or contexts in some portfolio entries.	Student performs targeted skills in a variety of settings or contexts in most portfolio entries.	Student performs targeted skills in an extensive variety of settings or contexts within and across all portfolio entries.	
Social Relationships	The student has no or limited social interactions during activities with others, both with and without disabilities, in portfolio entries	The student has some social interactions during activities with others, with and without disabilities, in some portfolio entries.	The student has sustained social interactions during activities with others, with and without disabilities, in most portfolio entries.	The student has varied, sustained social interactions during activities with others, with and without disabilities, in all portfolio entries.	
Self-Determination	The student makes no or limited choices in planning, monitoring, or evaluating own activities in the portfolio entries.	The student makes some choices in planning, monitoring, or evaluating own activities in some portfolio entries.	The student makes choices in planning, monitoring, or evaluating own activities in most portfolio entries.	The student consistently makes choices in planning, monitoring, or evaluating own activities within and across all portfolio entries.	
				PART II TOTAL *	

Portfolio Score Summary

(Scorer Use Only)

CONTENT AREA	PART I SCORE +	PART II TOTAL* =	TOTAL SCORE
Communication			
Reading			
Writing			
Mathematics			

Appendix I

Washington Alternate Assessment Portfolio Scoring Summary

Student Name*_____ *Portfolio Number

Part I: Progress on IEP Skills

(Progress on IEP skills scored separately for each content area entry.)

	1	2	3	4
Progress on IEP Skill linked to EALRs	Little or no progress on targeted skills linked to the EALRs in portfolio entry.	Clear progress on targeted skills linked to the EALRs in portfolio entry.	Attains goal for targeted IEP skills linked to the EALRs in portfolio entry.	Exceeds goal for targeted IEP skills linked to the EALRs in portfolio entry.

CONTENT AREA	PART I SCORE
Communication (Grade 4, 7, 10)	
Reading (Grade 4, 7, 10)	
Writing (Grade 4, 7, 10)	
Mathematics (Grade 4, 7, 10)	
Science (Grade 8 & 10)	

Part II: Student Generalization of Skills

(These dimensions are scored across the entire portfolio.)

Dimension	1	2	3	4	Scorer Use Only
Modifications and Adaptations	No or limited evidence that the student uses supports, modifications, adaptations or assistive technology in portfolio entries.	The student appropriately uses supports, modifications, adaptations or assistive technology in some portfolio entries.	The student appropriately uses supports, modifications, adaptations or assistive technology in most portfolio entries.	The student appropriately uses natural supports, modifications, adaptations or assistive technology within and across all portfolio entries.	
Settings and Contexts	Student participates in a limited number of settings or use of targeted skills unclear in portfolio entries.	Student performs targeted skills in some settings or contexts in some portfolio entries.	Student performs targeted skills in a variety of settings or contexts in most portfolio entries.	Student performs targeted skills in an extensive variety of settings or contexts within and across all portfolio entries.	
Social Relationships	The student has no or limited social interactions during activities with others, both with and without disabilities, in portfolio entries	The student has some social interactions during activities with others, with and without disabilities, in some portfolio entries.	The student has sustained social interactions during activities with others, with and without disabilities, in most portfolio entries.	The student has varied, sustained social interactions during activities with others, with and without disabilities, in all portfolio entries.	
Self-Determination	The student makes no or limited choices in planning, monitoring, or evaluating own activities in the portfolio entries.	The student makes some choices in planning, monitoring, or evaluating own activities in some portfolio entries.	The student makes choices in planning, monitoring, or evaluating own activities in most portfolio entries.	The student consistently makes choices in planning, monitoring, or evaluating own activities within and across all portfolio entries.	

Appendix J

Washington Alternate Assessment System Portfolio Academic Achievement Standard Descriptions

The academic achievement standards for students with significant disabilities who are participating in the Washington Alternate Assessment System (WAAS) portfolio are significantly different than the standards for students who participate in the Washington Assessment of Student Learning (WASL). The WAAS portfolio is based on the Essential Academic Learning Requirements (EALR) Extensions which allow the student to participate and progress in the general curriculum. Because the WAAS portfolio is based on the student's Individualized Education Program (IEP) goals in relation to the EALR Extensions, the specific assessment targets selected for the student may be the same for many content areas but may be different than for any other student. Additionally, these students have educational goals that may remain the same throughout their educational careers. Therefore, the following academic achievement standard descriptors apply for all grades and content areas.

Level 1 - Students performing at this level will be making little or no progress toward the goal for the targeted IEP skills linked to the EALRs. The student is unable to generalize the use of these targeted skills, using modifications and adaptations in any settings or contexts. The student cannot make choices in planning, monitoring or evaluating own performances. The student has no or limited social interactions with others during educational activities.

Level 2 - Students performing at this level will be making some progress toward the goal for the targeted IEP skills linked to the EALRs. The student is able to generalize the use of these targeted skills in some ways. The student may appropriately use modifications and adaptations in some settings and contexts or make choices in planning, monitoring or evaluating own performances. The student may have some social interactions with others during educational activities. The student is not able to generalize the targeted IEP skills in all of these ways.

Level 3 - Students performing at this level will be making clear progress or attaining the goal for the targeted IEP skills linked to the EALRs. The student is able to generalize the use of these targeted skills, appropriately using modifications and adaptations in a variety of settings and contexts while making choices in planning, monitoring or evaluating own performances. The student sustains some social interactions with others during educational activities.

Level 4 - Students performing at this level will be attaining or exceeding the goal for the targeted IEP skills linked to the EALRs. The student is able to generalize the use of these targeted skills, appropriately using natural supports, modifications or adaptations in an extensive variety of settings or contexts while consistently making choices in planning, monitoring or evaluating own performances. The student has sustained, varied social interactions with others during educational activities.

Appendix K

Standard Setting for the WAAS Portfolio Rating Form

Name _____

Portfolio # (xxx-xx-xx)	Content Area (C, R, W, M)	Scoring Pattern (x+x+x+x+x=xx)	Round 1 Level (1, 2, 3, 4)	Round 2 Level (1, 2, 3, 4)
		1-1-1-1-1		
		1-2-1-1-1		
		2-2-1-1-1		
		1-3-2-1-1		
		3-2-1-1-1		
		1-2-2-2-1		
		1-4-1-1-2		
		2-3-2-1-1		
		1-1-3-2-3		
		2-4-1-1-2		
		1-4-2-2-2		
		2-3-2-2-2		
		3-4-1-1-2		
		1-3-3-2-3		
		2-4-2-2-2		
		3-1-3-3-2		
		3-3-2-2-2		
		2-3-3-3-2		
		3-4-2-2-2		
		4-1-3-3-2		
		1-4-4-3-2		
		2-4-3-1-4		
		1-4-4-3-3		
		4-3-3-2-3		
		1-4-4-4-3		
		2-4-4-3-3		
		3-4-4-3-2		
		3-4-4-3-3		
		2-4-4-4-4		
		3-4-4-4-3		
		3-4-4-4-4		
		4-4-4-4-4		

Final Recommendations for Cut scores

Level 2 Total Score_____ Level 2 Total Score_____ Level 2 Total Score_____
Level 3 Total Score_____ Level 3 Total Score_____ Level 3 Total Score_____
Level 4 Total Score_____ Level 4 Total Score_____ Level 4 Total Score_____

Appendix L

Relationship Among Total Score, Level and Score Patterns

Total Score	Level	Pattern				
		Cont	M & A	S & C	S R	S D
20	4	4	4	4	4	4
19	4	3	4	4	4	4
18	4	3	4	4	4	3
18	3	2	4	4	4	4
17	4	3	4	4	3	3
16	4	3	4	4	3	2
16	3	2	4	4	3	3
16	2	1	4	4	4	3
15	3	4	3	3	2	3
15	2	1	4	4	3	3
14	3	2	4	3	1	4
14	2	1	4	4	3	2
13	3	4	1	3	3	2
13	3	3	4	2	2	2
13	3	2	3	3	3	2
12	3	3	3	2	2	2
12	3	3	1	3	3	2
12	3	2	4	2	2	2
12	2	1	3	3	2	3
11	2	3	4	1	1	2
11	2	2	3	2	2	2
11	2	1	4	2	2	2
10	2	2	4	1	1	2
10	2	1	1	3	2	3
9	2	2	3	2	1	1
9	2	1	4	1	1	2
8	2	1	2	2	2	1
8	2	3	2	1	1	1
8	2	1	3	2	1	1
7	1	2	2	1	1	1
6	1	1	2	1	1	1
5	1	1	1	1	1	1

Cont – Content Area – Communication, Reading, Writing, Mathematics, Science

M & A - Modifications and Adaptations

S & C - Settings and Contexts

SR - Social Relationships

SD - Self-Determination